Single Top Production at the Tevatron

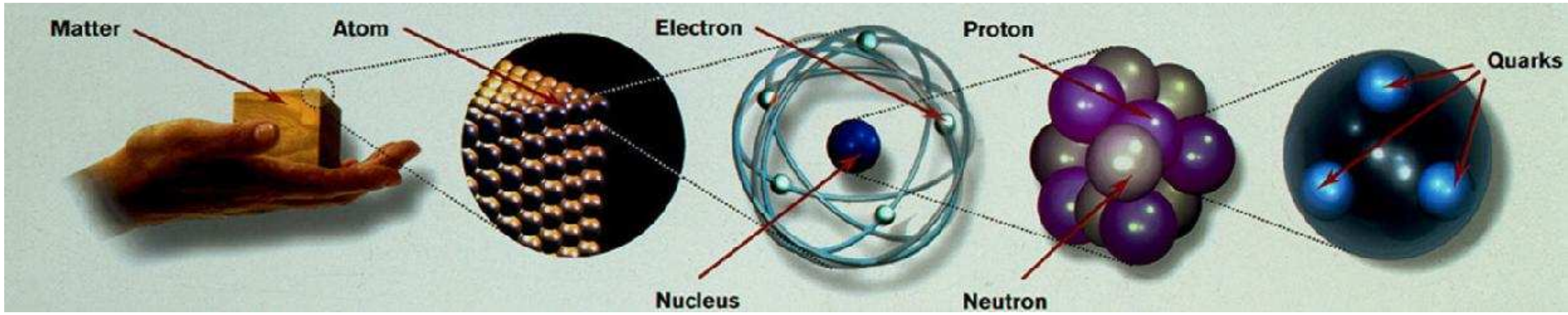
Daniel Wicke
(Bergische Universität Wuppertal)

Outline

- Introduction
- DØ Cross Section
- CDF Results
- DØ $|V_{tb}|$
- Conclusions

Introduction and Motivation

The Standard Model of Elementary Particle Physics



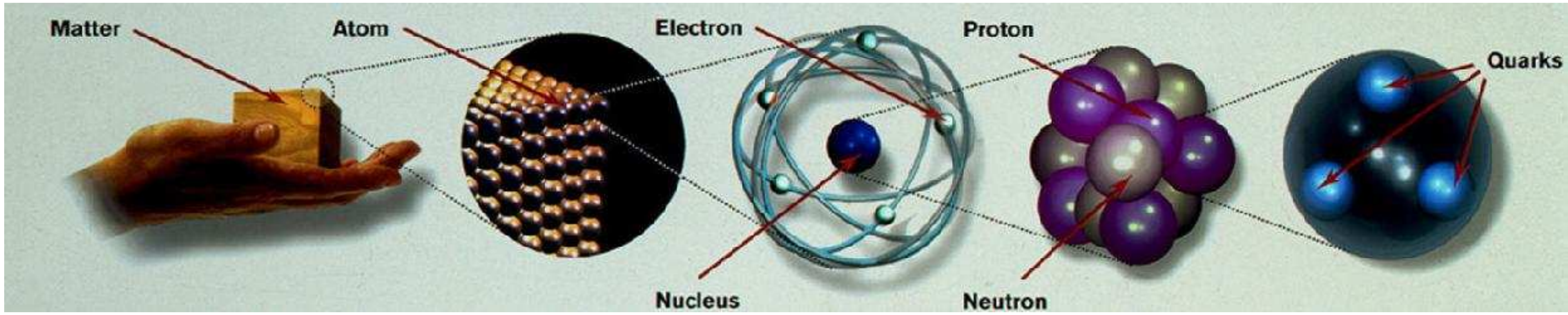
Matter

e	electron
ν_e	electron neutrino
u	up-quark
d	down-quark

Forces

Electromagnetism: Photon γ, Weak force: $Z, W^{\pm}$, Strong force: Gluon g

The Standard Model of Elementary Particle Physics



Three families of matter

e	electron	μ	muon	τ	tauon
ν_e	electron neutrino	ν_μ	muon neutrino	ν_τ	tau neutrino
u	up-quark	c	charm-quark	t	top-quark
d	down-quark	s	strange-quark	b	bottom-quark

Forces

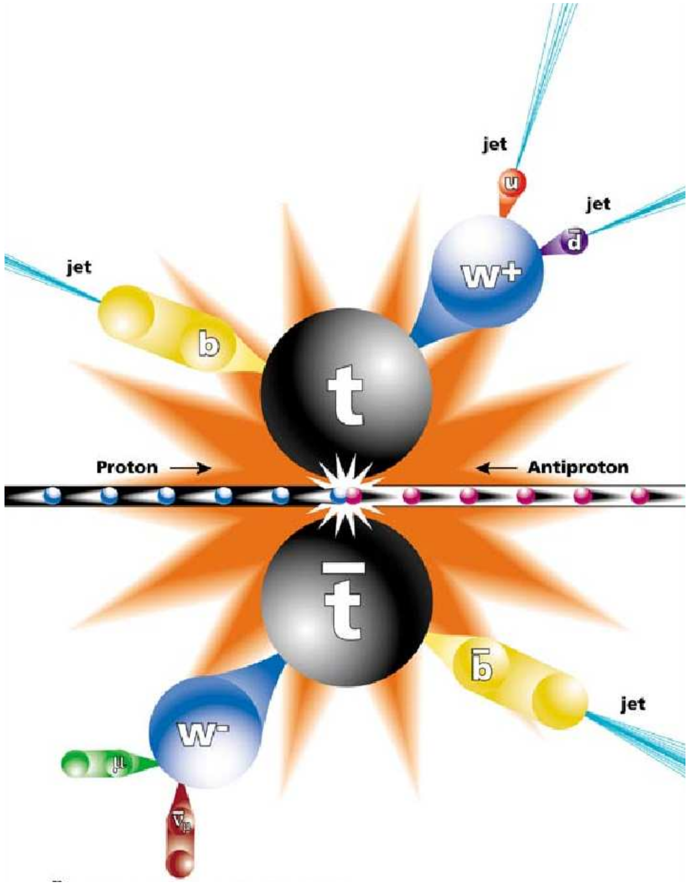
Electromagnetism: Photon γ, Weak force: $Z, W^\pm$, Strong force: Gluon g

The Top Quark

- Discovered by CDF and DØ in 1995.

- Completes set of quarks in SM.
- Quantum numbers as for up-type quarks.
- Only its mass is a free parameter.

- Production and decay properties fully defined in Standard Model.



Only few of its predicted properties verified

Is the Top Quark special?

Yes, it is! It is ...

- more than 30 time heavier than the second heaviest elementary Fermion.
  Its mass is surprisingly close to electro-weak scale.
- the only bare quark (i.e. decays before it hadronises)
- the last of the predicted quarks.

No, it isn't! It has ...

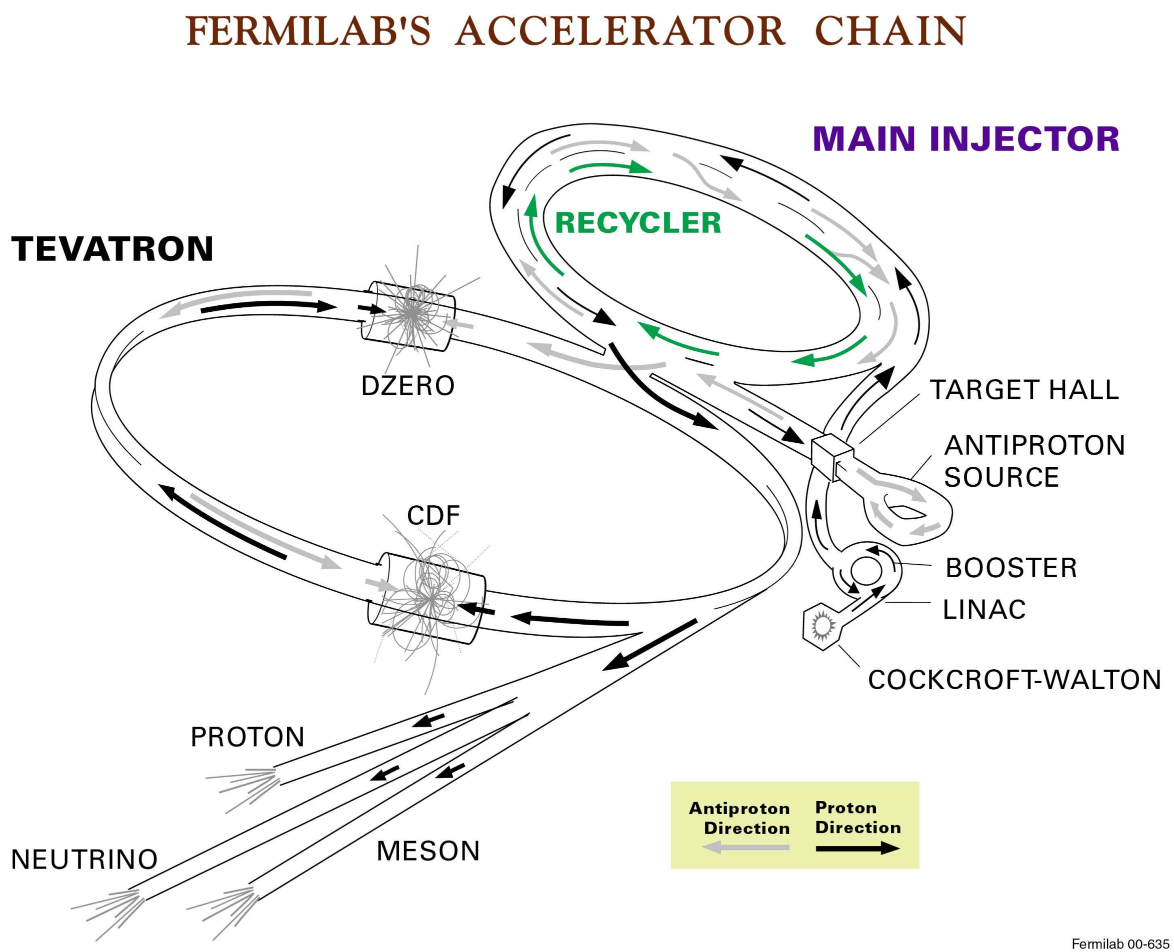
- the same electrical charge as u and c-quarks.
- the same weak couplings as the other quarks.
- the same colour charge as the other quarks. *Really??*

Maybe ...

- it has different couplings ($\Rightarrow$ new physics)
- it isn't last quark ($\Rightarrow$ new physics) *very interesting!*

The $p\bar{p}$ Accelerator Tevatron

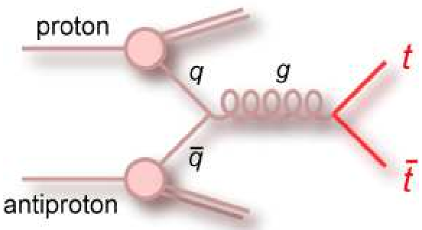
- Circumfence 7 km.
- $p\bar{p}$ collisions
- Run I (1987-1995)
- Run II (since 2001)
  Collision energy 2 TeV
- 2 experiments,
  CDF and DØ,
  record events.



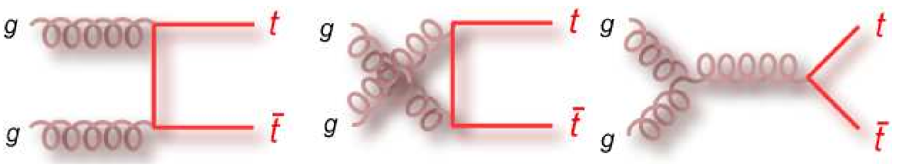
The Tevatron

Top Quark Production at the Tevatron

Strong top production

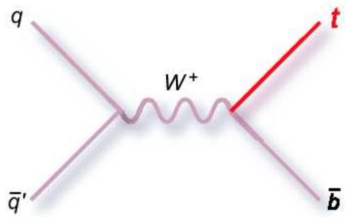


$$\begin{array}{ll} 85\% & q\bar{q} \rightarrow t\bar{t} \\ 15\% & gg \rightarrow t\bar{t} \end{array}$$



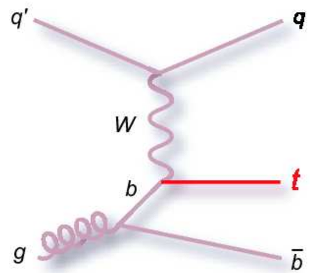
- $\sigma(t\bar{t}) = 6.77 \pm 0.42$ pb

Weak top production



Drell-Yan

s-channel



W-g fusion

t-channel

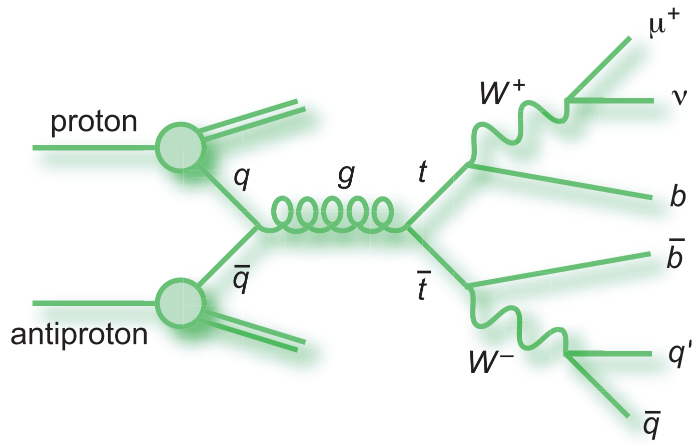
- $\sigma(t) = 2.9 \pm 0.3$ pb

For integrated luminosity of $\sim 1\,\mathrm{fb}^{-1}$

around 7000 top pairs and 3000 single tops expected.

Top Quark Decay

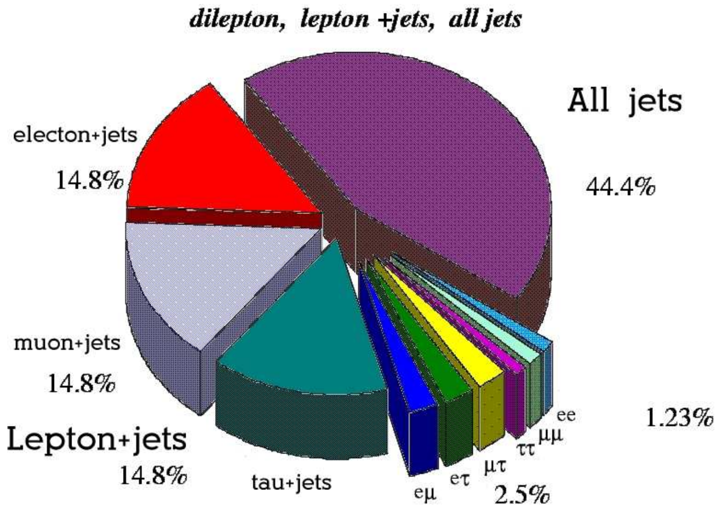
Top quarks decay to bW (nearly) 100%.



Pair Production Signatures

Decay modes are defined by W-decays:

- Dilepton $(2b + 2l + 2\nu)$
- Lepton+jets $(2b + 2q + l\nu)$
- Alljets $(2b + 4q)$



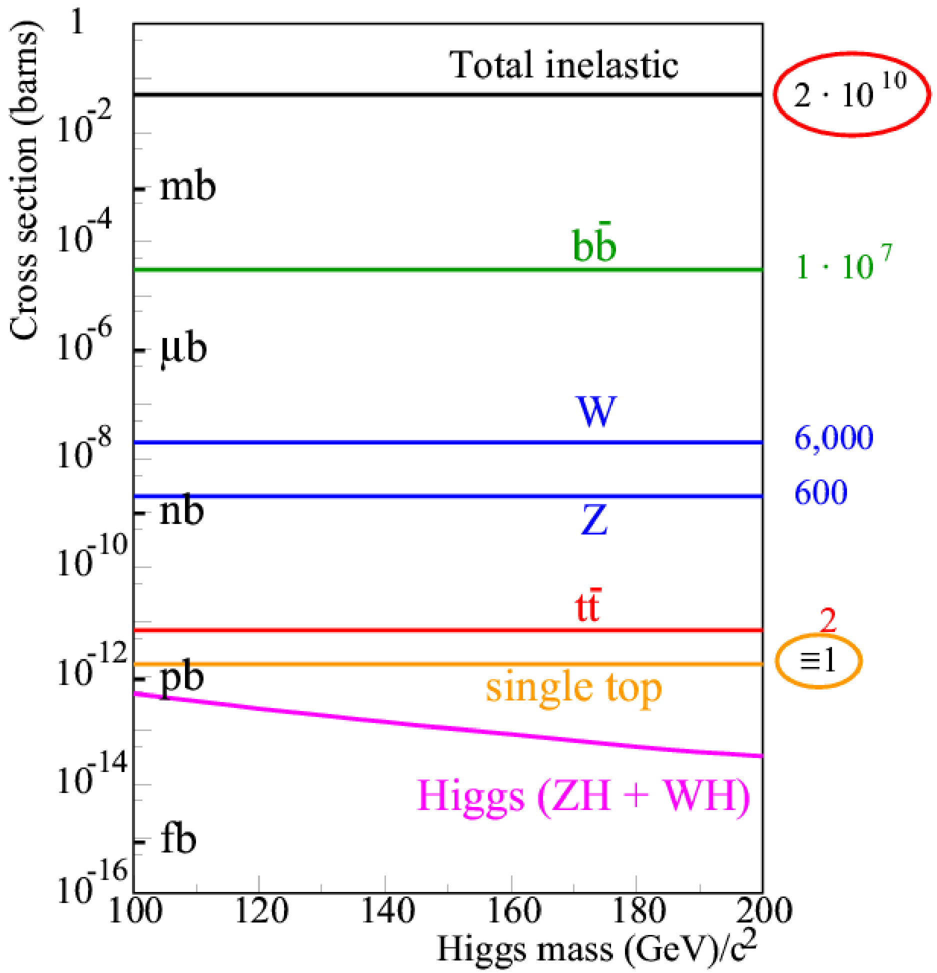
Single Top Signatures

Defined by W-decays and channel;
e.g. leptonic decay:

- s-channel $(2b + l + \nu)$
- t-channel $(b + j + l + \nu)$

Backgrounds

$$\sigma_{t\bar{t}} = \frac{N - B}{\varepsilon \mathcal{L} \cdot \mathrm{BR}}$$



Physics background

- Multijet ($q\bar{q}$ or gg + gluon rad.)
- W+jets
- Z+jets

Instrumental background

- Physics object misidentification
- Mismeasurement of energies

Small, but amplified by cross-section.

The DØ Detector

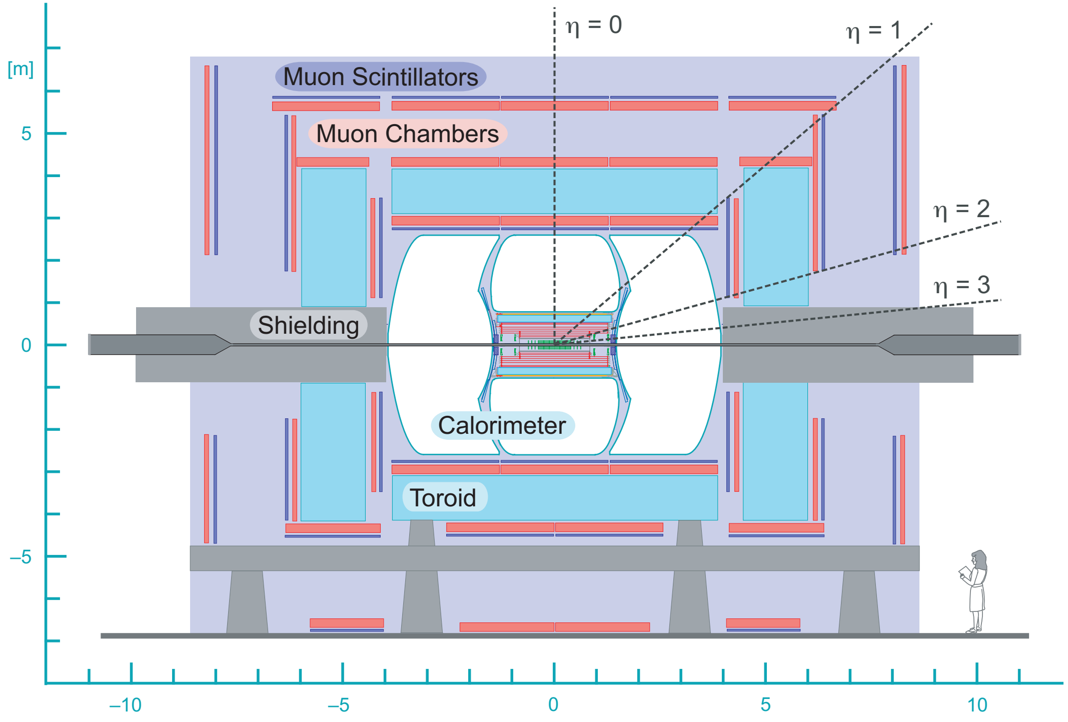
A 4π general purpose detector:

- Tracking in 2T solenoid
  - Silicon microstrip
  - Scintillating fiber tracker
- Calorimetry
  - Uranium/liquid argon
- Muon spectrometer
  - 3 layers of drift tubes
  - Toroidal magnetic field (1.9T between inner 2 layers)

Dimensions: $12 \times 12 \times 20\mathrm{m}^3$

Note: Polarangle θ against beam axis

Pseudorapidity $\eta = -\ln \tan \theta/2$



Reconstructed Physics Objects

Muon

Track in Muon chambers (outside the calorimeters)

Electron

Energy deposition only in the innermost ('em') calorimeter part.

Jets (sign of quarks or gluons)

Accumulations of energy deposited in the 'hadron' calorimeters.
CDF and DØ usually use Cone jet algorithms.

Missing Transvers Energy, $\not{E}_T$ (sign of neutrinos)

Negative sum of all energy measured transverse to beam directions

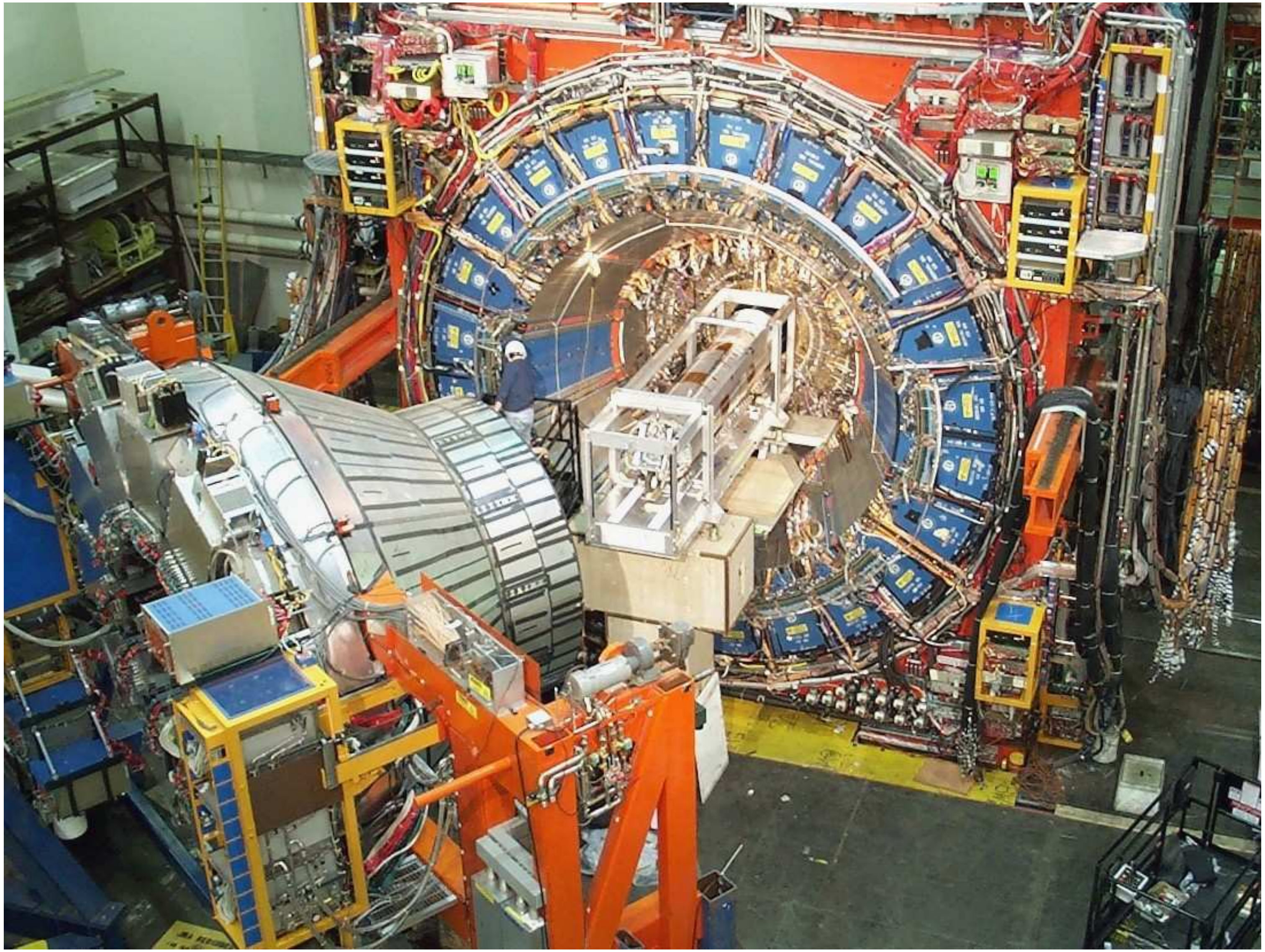
B-Tag (sign of b-quarks)

Long lifetime of B-Hadrons lead to secondary vertices, detected with tracking.

DØ

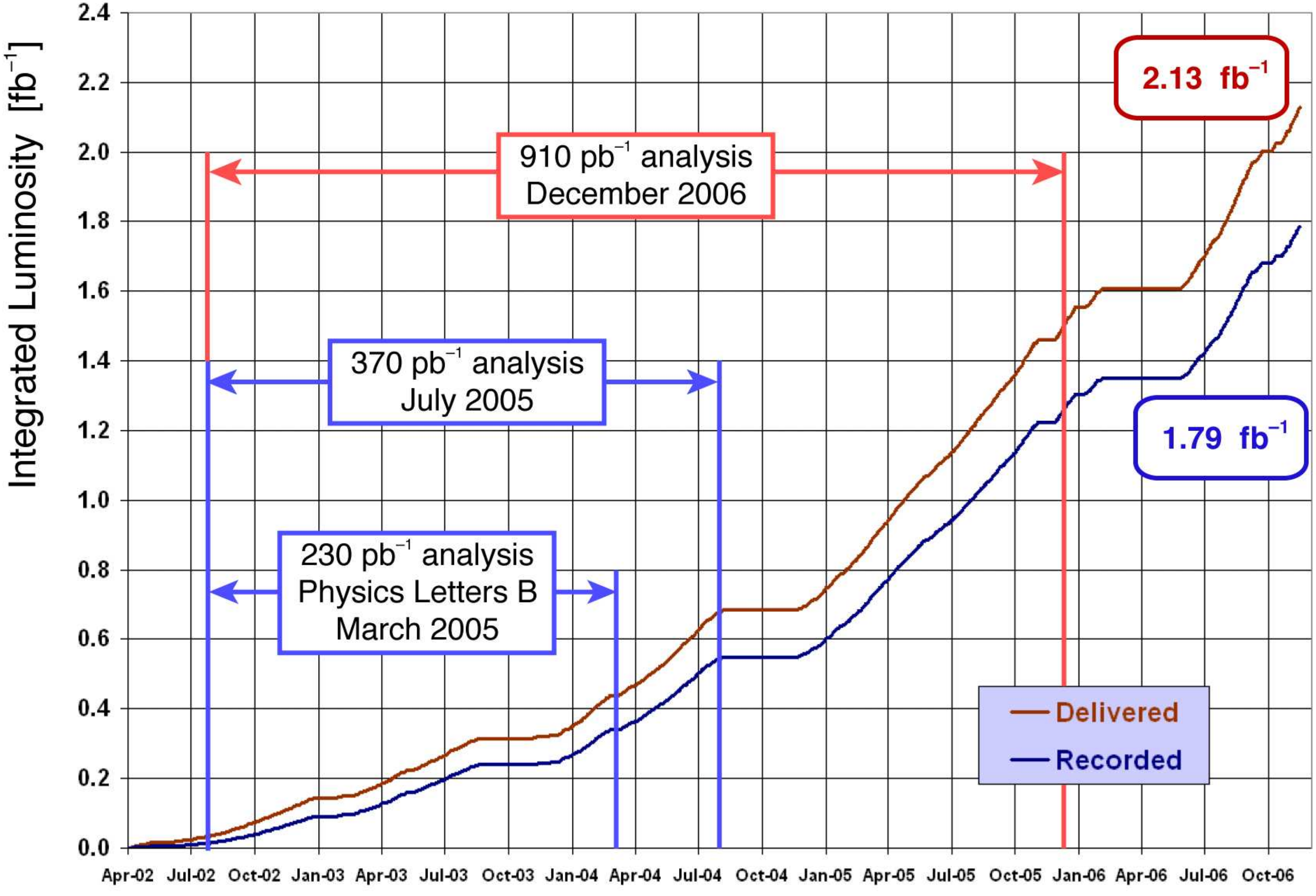
- International collaboration
- 86 institutes, 670 physicists
- 19 countries from 4 continents

The CDF Detector



DØ RunII Integrated Luminosity

Apr 2002 – Dec 2006



Many thanks to the Accelerator Division

DØ Cross Section

Signal signatures

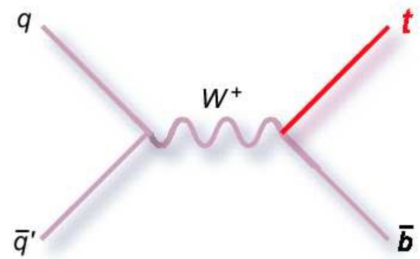
(Consider only leptonic decay $t \rightarrow bW \rightarrow b + l\nu$.)

s-channel

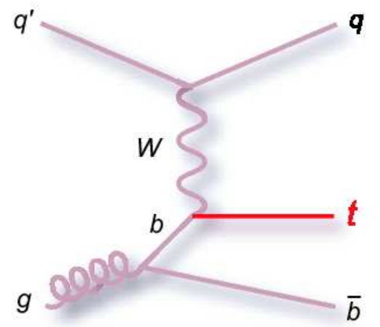
- 2 b-quarks $\rightarrow$ 2 b-tagged jets.
- 1 lepton $\rightarrow$ electron or muon.
- 1 neutrino $\rightarrow$ missing transverse energy

t-channel

- 1 b-quark and 1 light quark $\rightarrow$ 2 jets (1 b-tagged).
- 1 lepton $\rightarrow$ electron or muon.
- 1 neutrino $\rightarrow$ missing transverse energy

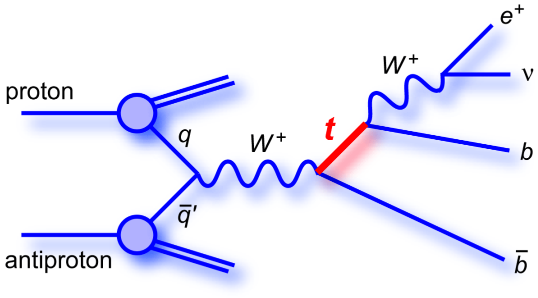


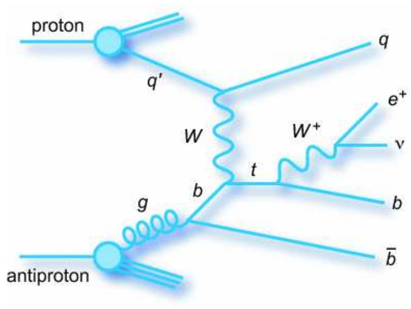
s-channel



t-channel

Event Selection



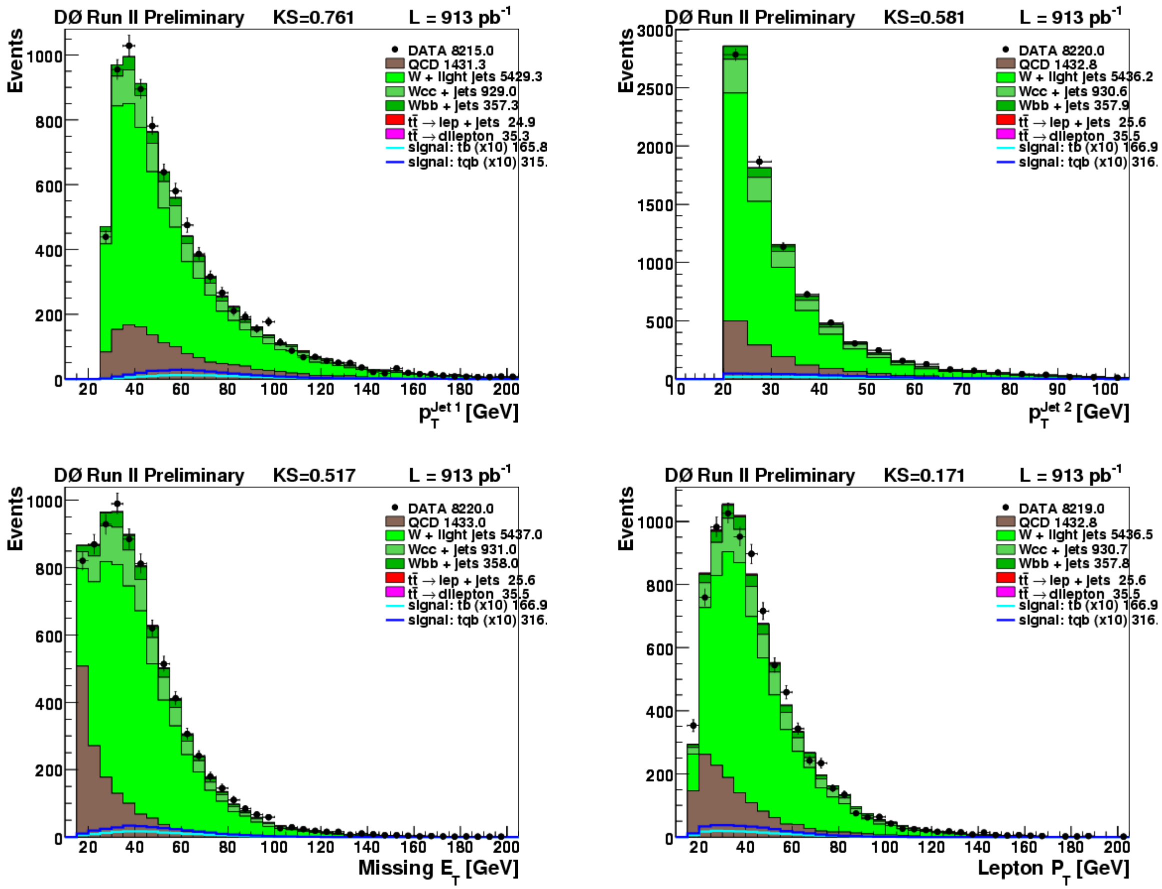


Signature

- isolated lepton
- $\not\!\!E_T$
- 2 – 4 jets
- at least 1 b-jet

- Only one tight and no other loose lepton
  - electron: $p_T > 15$ GeV and $|\eta_{det}| < 1.1$
  - muon: $p_T > 18$ GeV and $|\eta_{det}| < 2$
- $15 < \not\!\!E_T < 200$ GeV
- 2 – 4 jets with $p_T > 15$ GeV and $|\eta_{det}| < 3.4$
  - Leading jet with $p_T > 25$ GeV and $|\eta_{det}| < 2.5$
  - Second leading jet $p_T > 20$ GeV

MC-Data Agreement – before tagging



- Normalize W+multijet to data before tagging
- Checked 90 variables, 3 jet multiplicities, 1-2 tags, electron + muon
- Shown: electron, 2 jets, before tagging
- Good description of data

Event yields

Source	Event Yields in 0.9 fb^{-1} Data Electron+muon, 1tag+2tags combined 2 jets	3 jets	4 jets
tb	16 ± 3	8 ± 2	2 ± 1
tqb	20 ± 4	12 ± 3	4 ± 1
$t\bar{t} \rightarrow ll$	39 ± 9	32 ± 7	11 ± 3
$t\bar{t} \rightarrow l$+jets	20 ± 5	103 ± 25	143 ± 33
W+$b\bar{b}$	261 ± 55	120 ± 24	35 ± 7
W+$c\bar{c}$	151 ± 31	85 ± 17	23 ± 5
W+jj	119 ± 25	43 ± 9	12 ± 2
Multijets	95 ± 19	77 ± 15	29 ± 6
Total background	686 ± 131	460 ± 75	253 ± 42
Data	697	455	246

Signal to Background

Percentage of single top *tb+tqb* selected events and S:B ratio (white squares = no plans to analyze)

Electron + Muon	1 jet	2 jets	3 jets	4 jets	≥ 5 jets
0 tags	10% 1 : 3,200	25% 1 : 390	12% 1 : 300	3% 1 : 270	1% 1 : 230
1 tag	6% 1 : 100	21% 1 : 20	11% 1 : 25	3% 1 : 40	1% 1 : 53
2 tags		3% 1 : 11	2% 1 : 15	1% 1 : 38	0% 1 : 43

Need multivariate analyses to disentagle signal and background

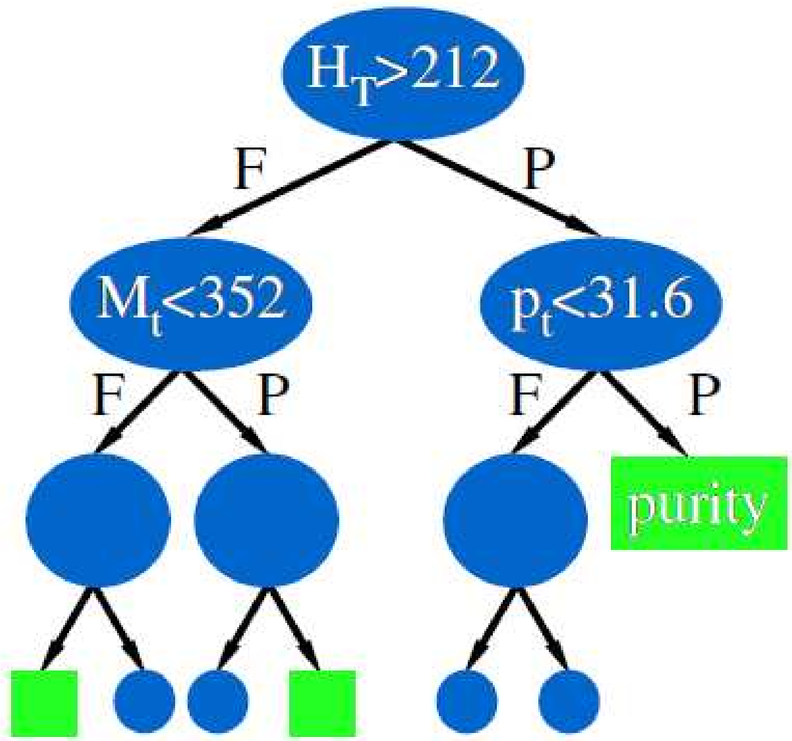
DØ Multivariate Analyses

- Decision Trees
- Testing Methods
- Matrix Element
- Bayesian Neural Networks
- Sensitivities and Results

Decision Trees

Training

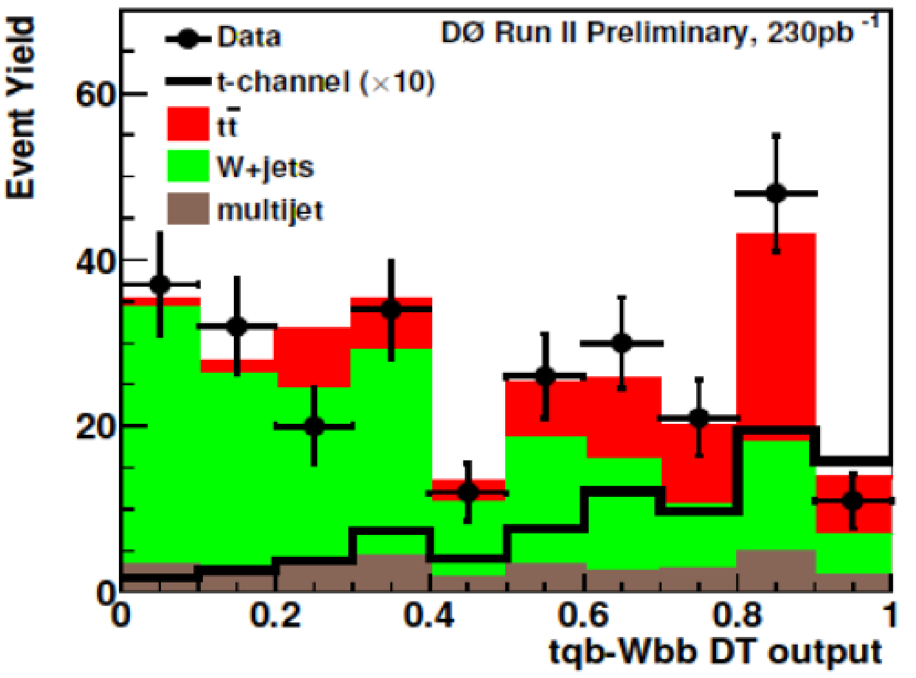
- Start with all events (first node)
- For each variable, find the splitting value with best separation between children (best cut).
- Select best variable and cut: produce Failed and Passed branches
- Repeat recursively on each node
- Stop when improvement stops or when too few events left. Terminal node = leaf.



Training performed with 49 observables on 12 subsamples

Measure and Apply

- Take trained tree and run on independent simulated sample, determine purities.
- Apply to Data:
  Tree output is leaf purity.
- Should see enhanced separation (signal right, background left)
- Could cut on output and measure, or use whole distribution to measure.



Boosting

Boosting

- Recent technique to improve performance of a weak classifier
- Recently used on DTs by GLAST and MiniBooNE
- Basic principal on DT:
  - train a tree T_k
  - $T_{k+1} = \text{modify}(T_k)$

AdaBoost algorithm

- Adaptive boosting
- Check which events are misclassified by T_k
- Derive tree weight α_k
- Increase weight of misclassified events
- Train again to build T_{k+1}
- Boosted result of event i: $T(i) = \sum_{n=1}^{N_{\text{tree}}} \alpha_k T_k(i)$

- Averaging dilutes piecewise nature of DT
- Usually improves performance

DØ uses 20 boosting cyles

Measuring the Cross Section

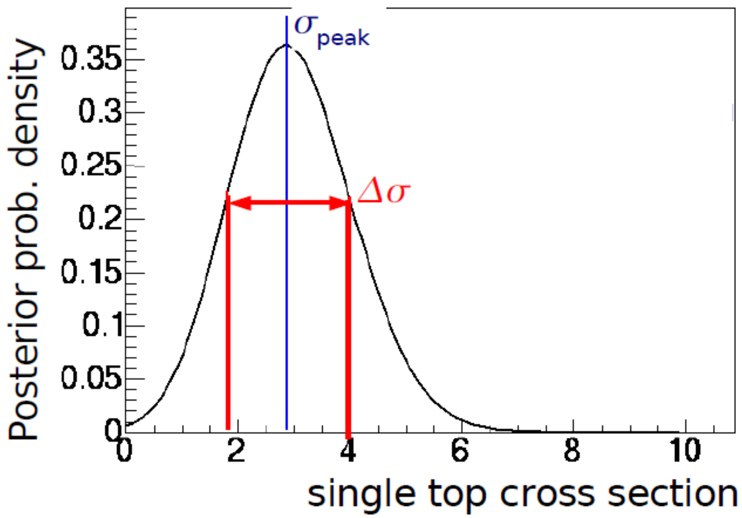
Probability to observe data distribution D, expecting $y = \underbrace{\alpha\,\mathcal{L}}_{a}\sigma + \sum_{s=1}^{N} b_s$

$$P(D|y) \equiv P(D|\sigma, a, b) = \prod_{i=1}^{n_{\text{bins}}} P(D_i|y_i)$$

The cross section is obtained

$$\text{Post}(\sigma|D) \equiv P(\sigma|D)$$

$$\propto \int_a \int_b P(D|\sigma, a, b)\text{Prior}(\sigma)\text{Prior}(a, b)$$



- Bayesian posterior probability density
- Shape and normalization systematics treated as nuisance parameters
- Correlations between uncertainties properly accounted for
- Flat prior in signal cross section

Ensemble Testing

- To verify that all of this machinery is working properly we test with many sets of **pseudo-data**.

- Wonderful tool to test analysis methods! Run DØ experiment 1000s of times!

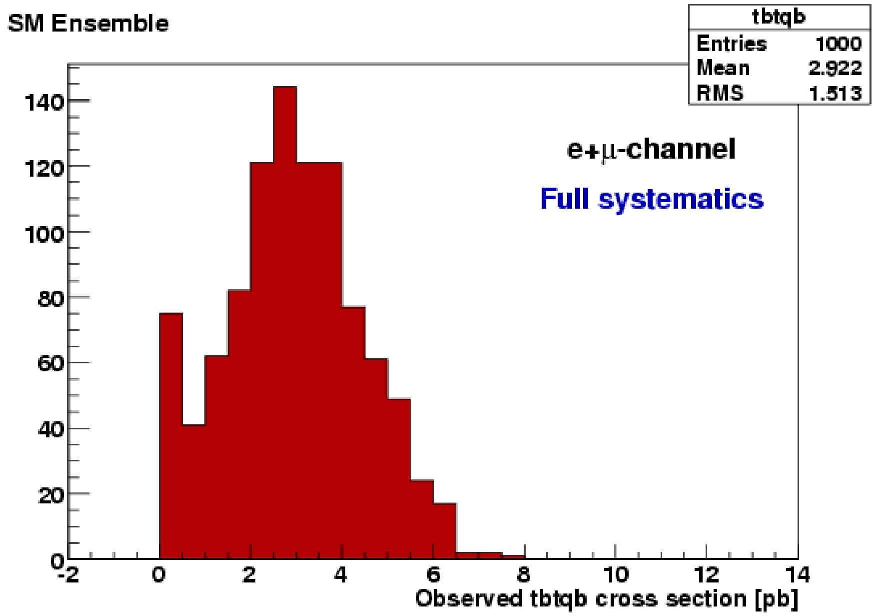
- Generated ensembles include:

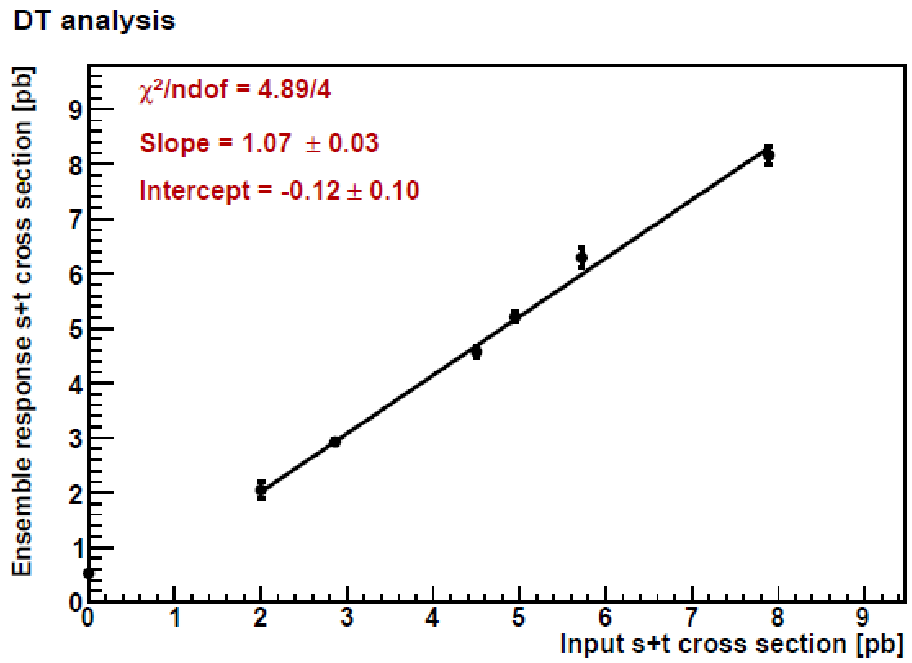
  1. 0-signal ensemble ($\sigma_{s+t} = 0$ pb)
  2. SM ensemble ($\sigma_{s+t} = 2.9$ pb)
  3. "Mystery" ensembles to test analyzers ($\sigma_{s+t} =??$ pb)
  4. Ensembles at measured cross section ($\sigma_{s+t} =$measured)
  5. A high luminosity ensemble

- Each analysis tests linearity of "response" to single top.

Decision Trees - Ensembles

- SM input is returned by DTs
- "Mystery" ensembles are unraveled by the DTs
- Linear response is achieved





Matrix Element

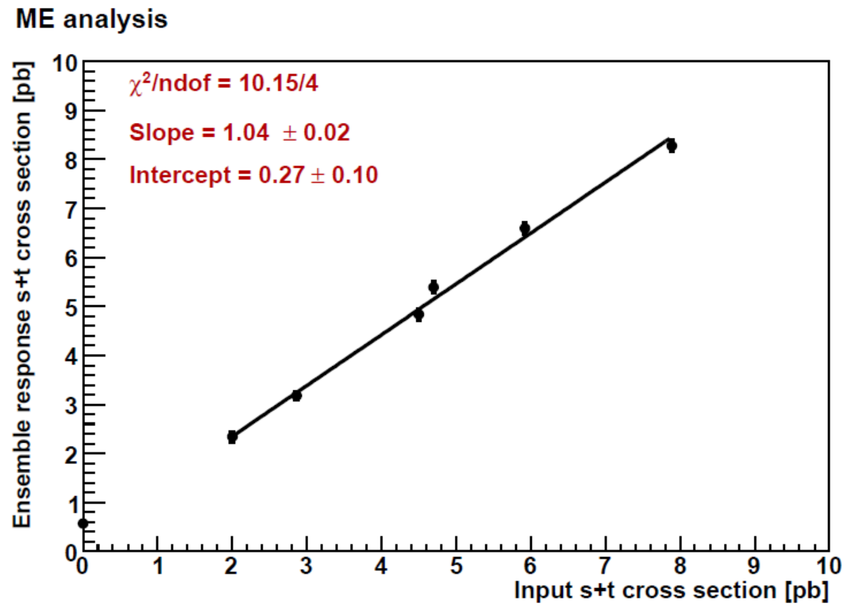
A matrix elements analysis takes a very different approach:

- Use the 4-vectors of all reconstructed leptons and jets
- Use matrix elements of main signal and background diagrams to compute an event probability density for signal and background hypotheses.
- Goal: calculate a discriminant:

$$D_s(\vec{x}) = P(S|\vec{x}) = \frac{P_{\mathrm{Sig}}(\vec{x})}{P_{\mathrm{Sig}}(\vec{x}) + P_{\mathrm{Bkg}}(\vec{x})}$$

- Define P_{Sig} as properly normalized differential cross section

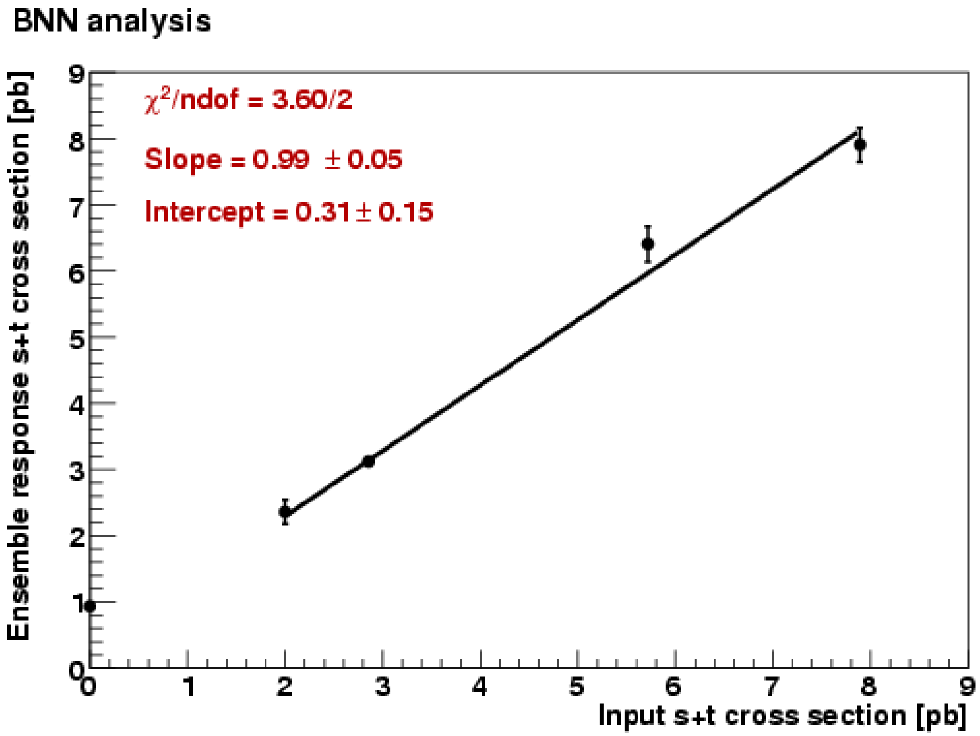
$$P_{\mathrm{Sig}}(\vec{x}) = \frac{1}{\sigma_S} \mathrm{d}\sigma_S(\vec{x})\sigma_S = \int \mathrm{d}\sigma_S(\vec{x})$$



Bayesian Neural Network

A special sort of neural network

- Determine posterior density of all possible weights
- Average over many networks.
- Less prone to over training.



Significances

Determination

Use 0-signal ensemble to determine sensitivity and significance:

Expected Significance (p-value)

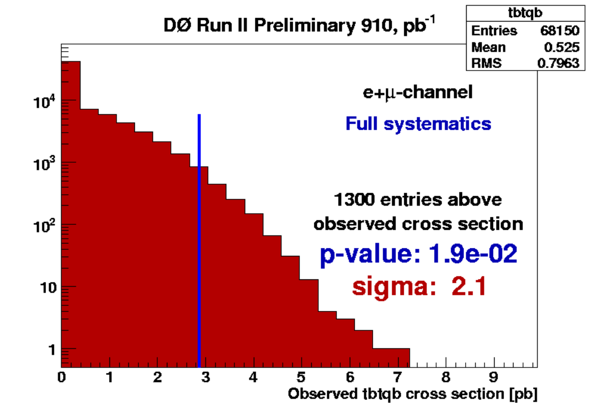
Fraction of 0-signal pseudo-datasets in which at least 2.9 pb is measured.

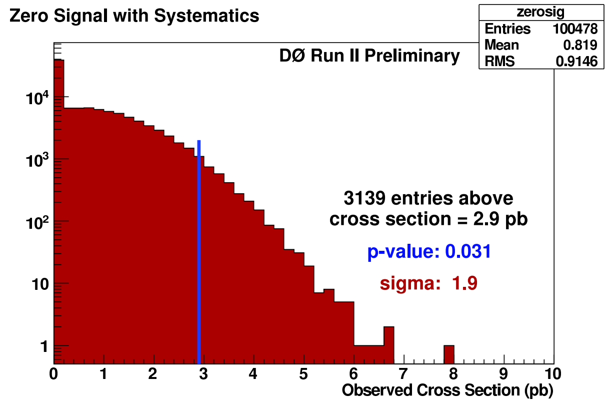
Observed Significance (p-value)

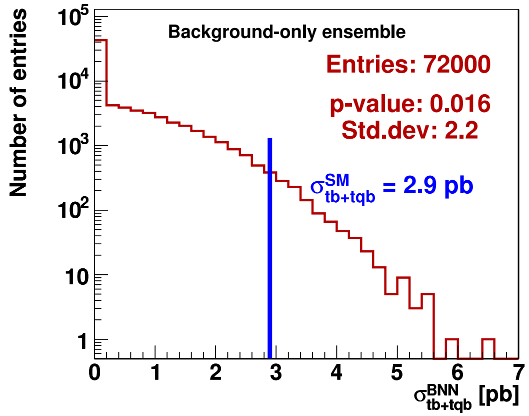
Fraction of 0-signal pseudo-datasets
in which at least the measured cross section is measured.

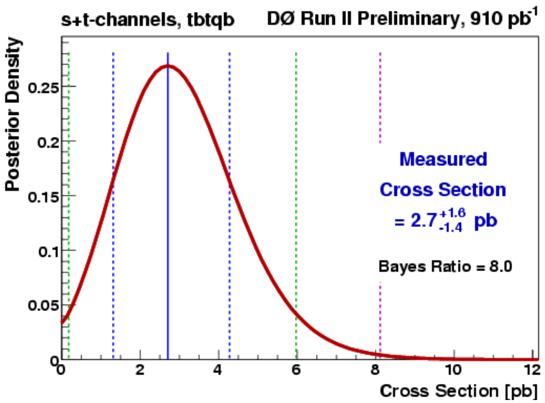
Expected Significances ($s + t$-channel)

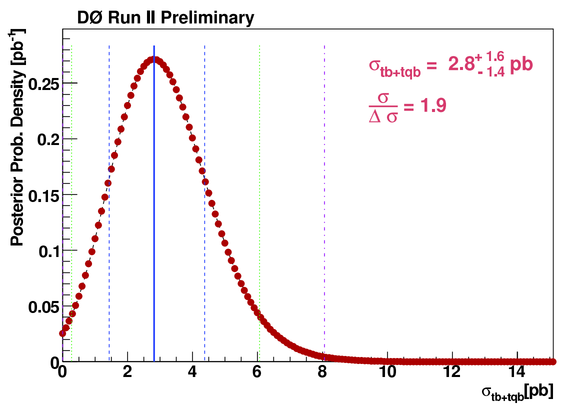
Decision Trees	Matrix Elements	Bayesian NN
p-value **1.8%**	p-value **3.1%** (was 3.7%)	p-value **1.6%** (was 9.7%)

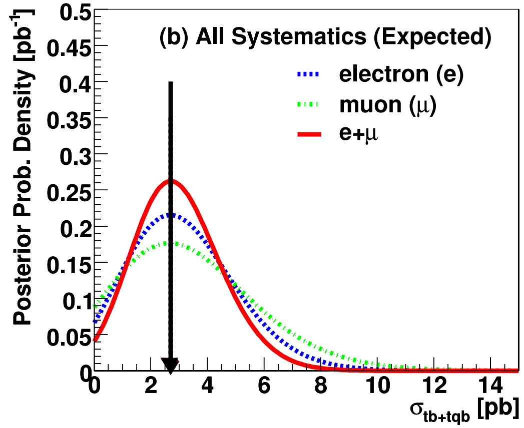






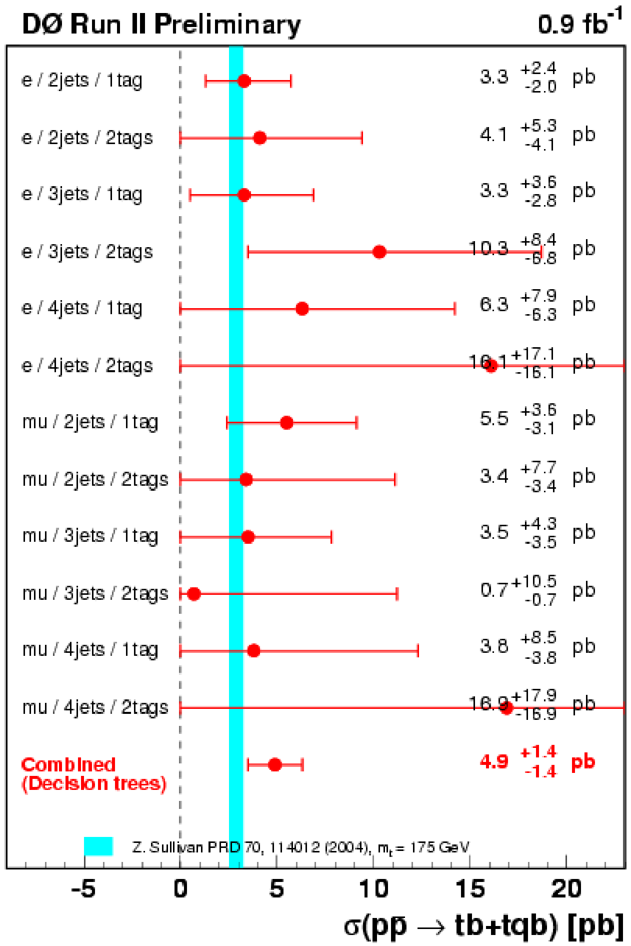


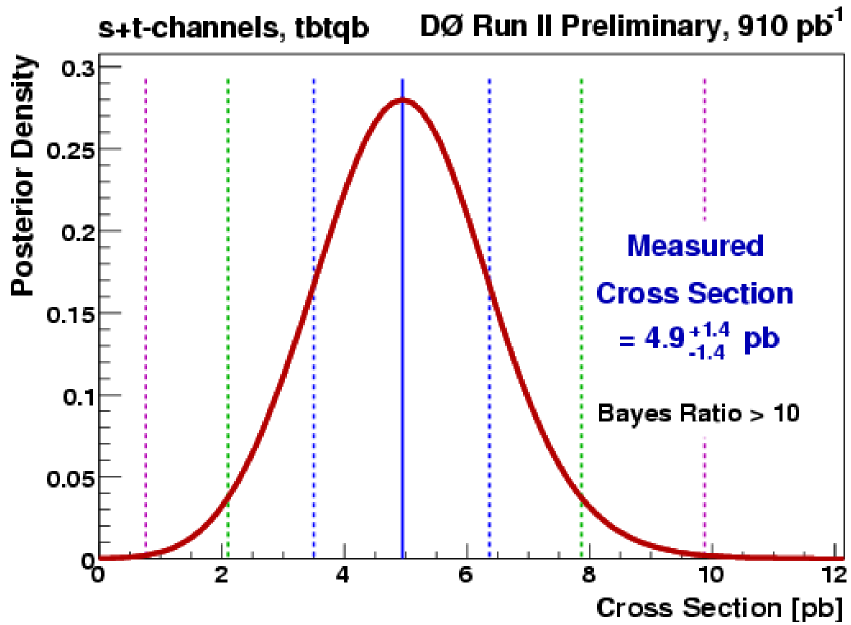




Observed Results

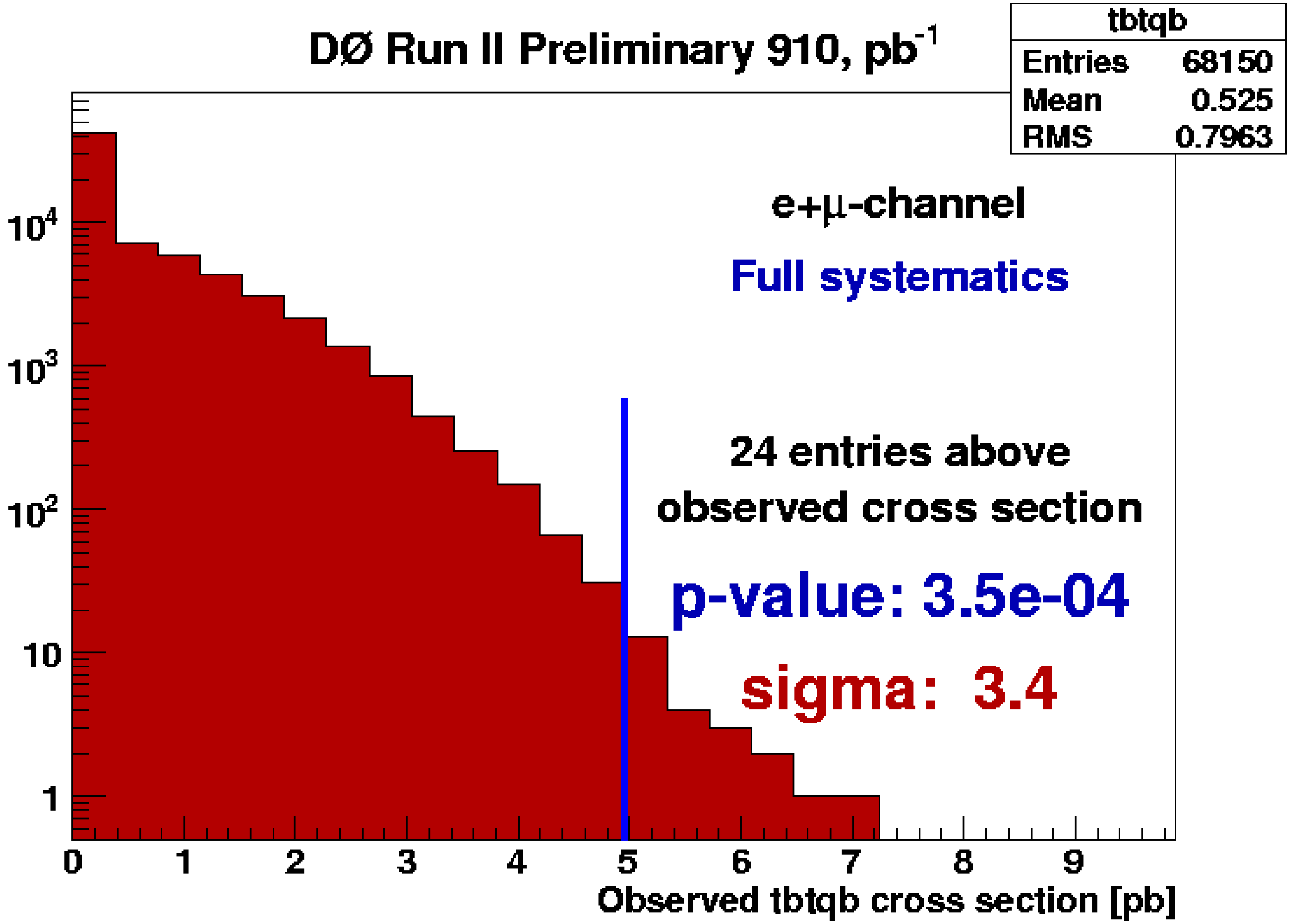
Decision Trees Cross Section (announced Dec. 2006)





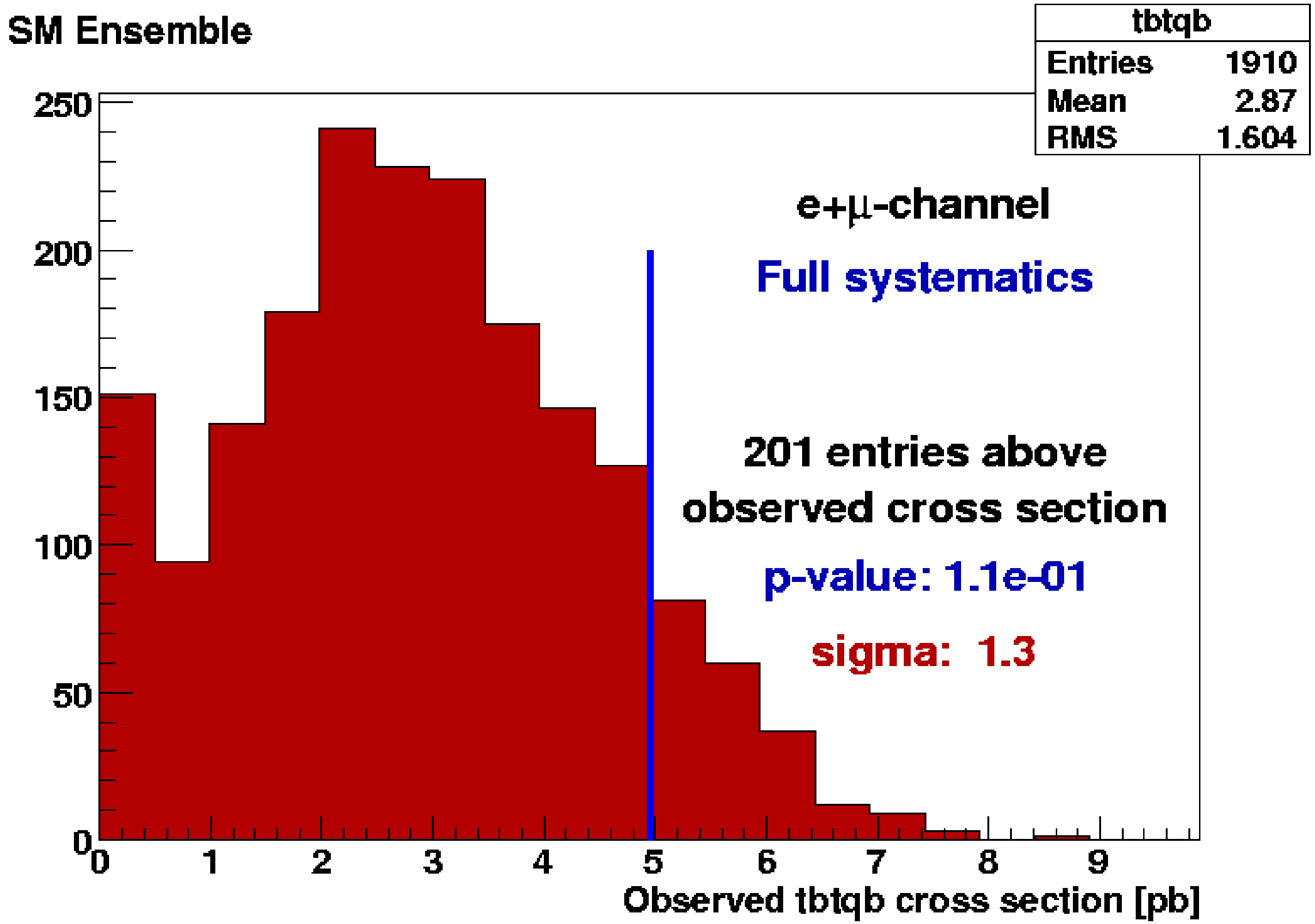
$$\sigma_{t+s} = 4.9 \pm 1.4\,\text{pb}$$

Decision Trees Observed Significance

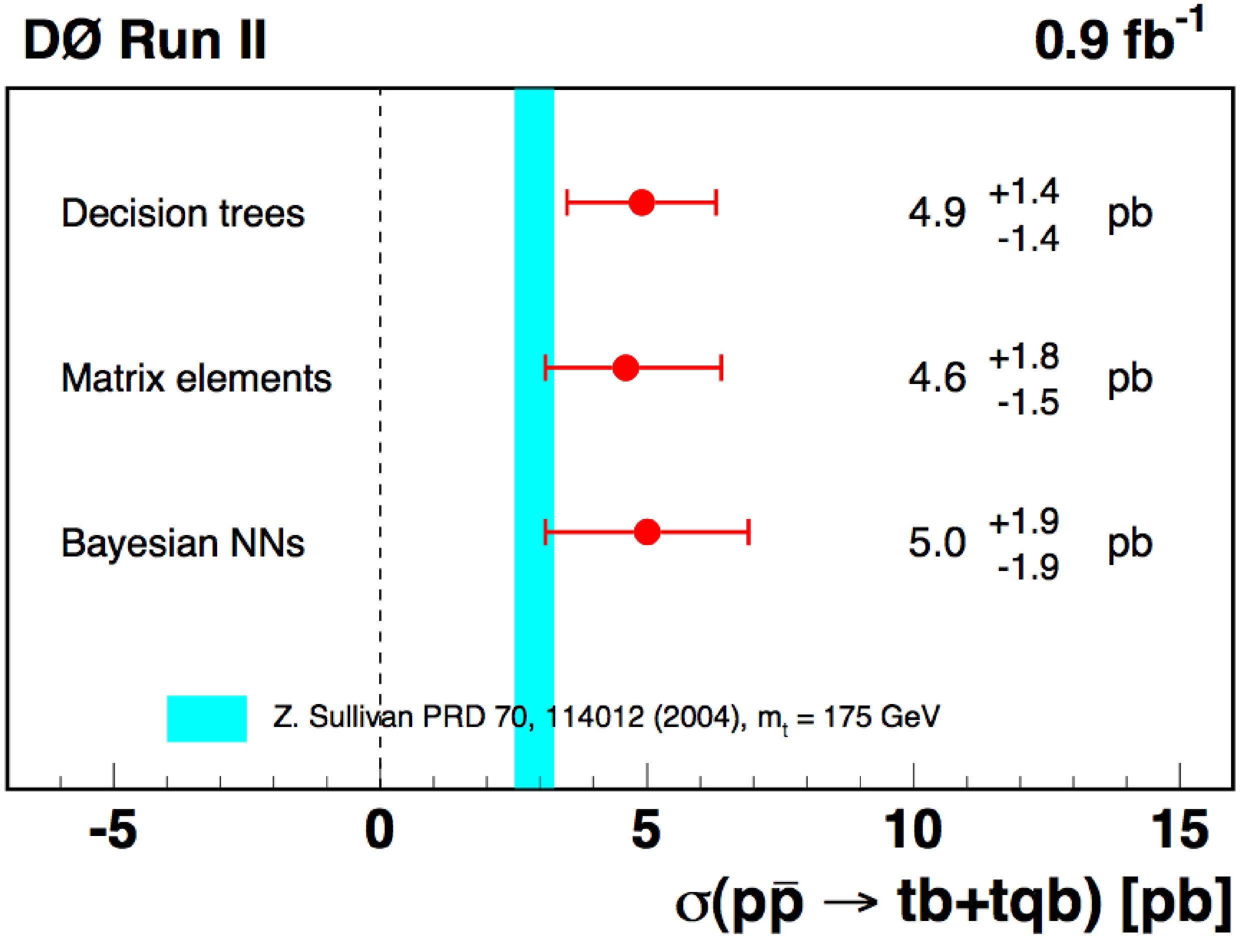


A 3.4 σ excess!!

SM Consistency



All Methods (as of Dec. 2006)



3.4σ

2.9σ

2.4σ

First evidence from DT analysis, consitent results with other methods

Correlations between methods

Is there room for improvement by combining the results?

Estimate correlations with ensemble tests:

	DT	ME	BNN
DT	100%	64%	66%
ME		100%	59%
BNN			100%

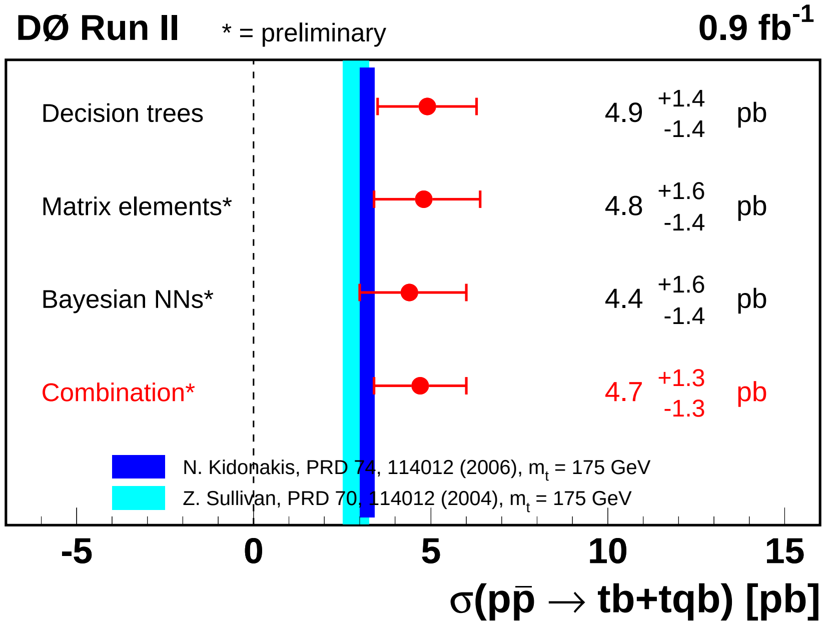
Yes!

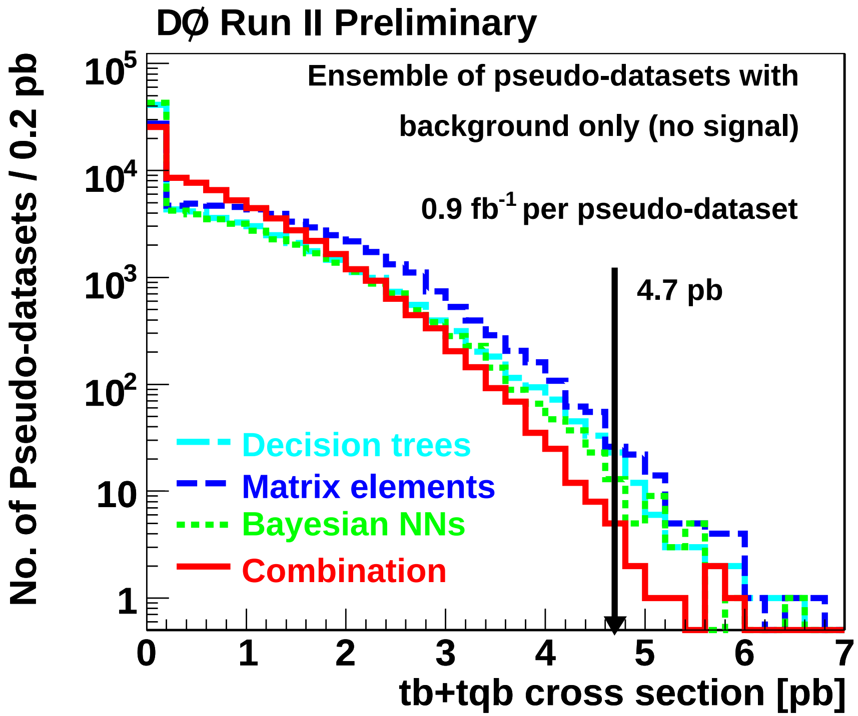
Average with Best Linear Unbiased Estimator (BLUE)

Weighted mean with weights from covariance matrix.

Combined Results

(as of June 2007)





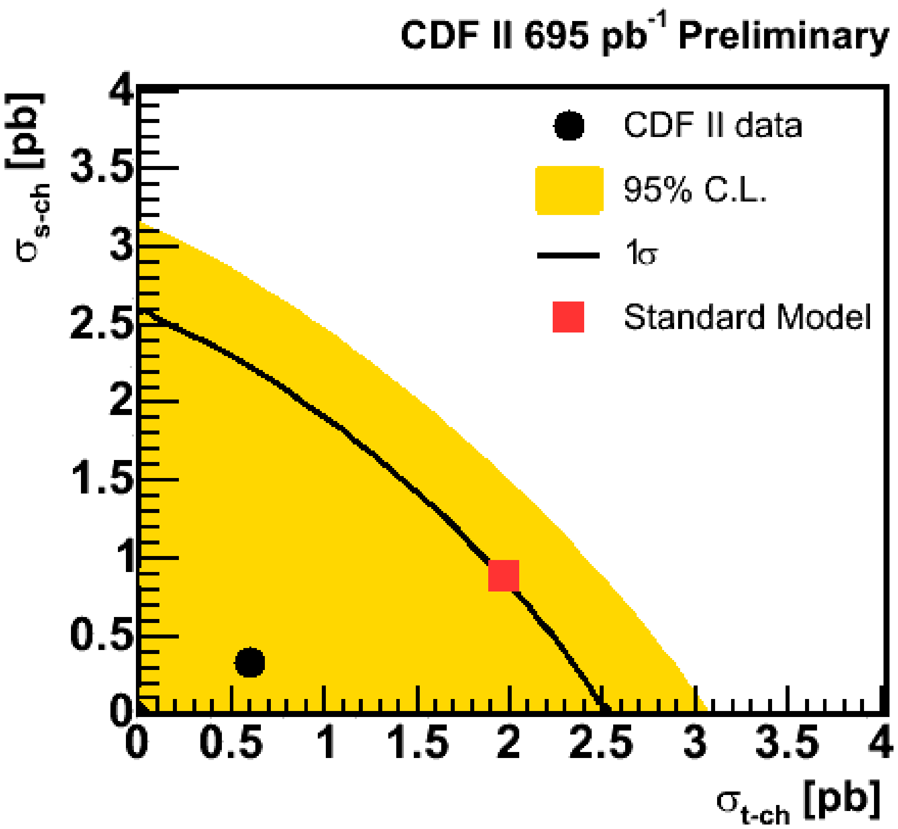
$$\sigma_{s+t} = 4.7 \pm 1.3\,\text{pb}$$

Probability for no signal hypothesis 0.014%, i.e. 3.6σ significance for deviation.

CDF Results

- Neural Network Technique
- Likelihood Function
- Matrix Element

Neural Network Analysis



Best fit Separate Search:

$$\sigma_s = 0.3^{+2.3}_{-0.3}\,\text{pb}$$

$$\sigma_t = 0.6^{+1.9}_{-0.6}\,\text{pb}$$

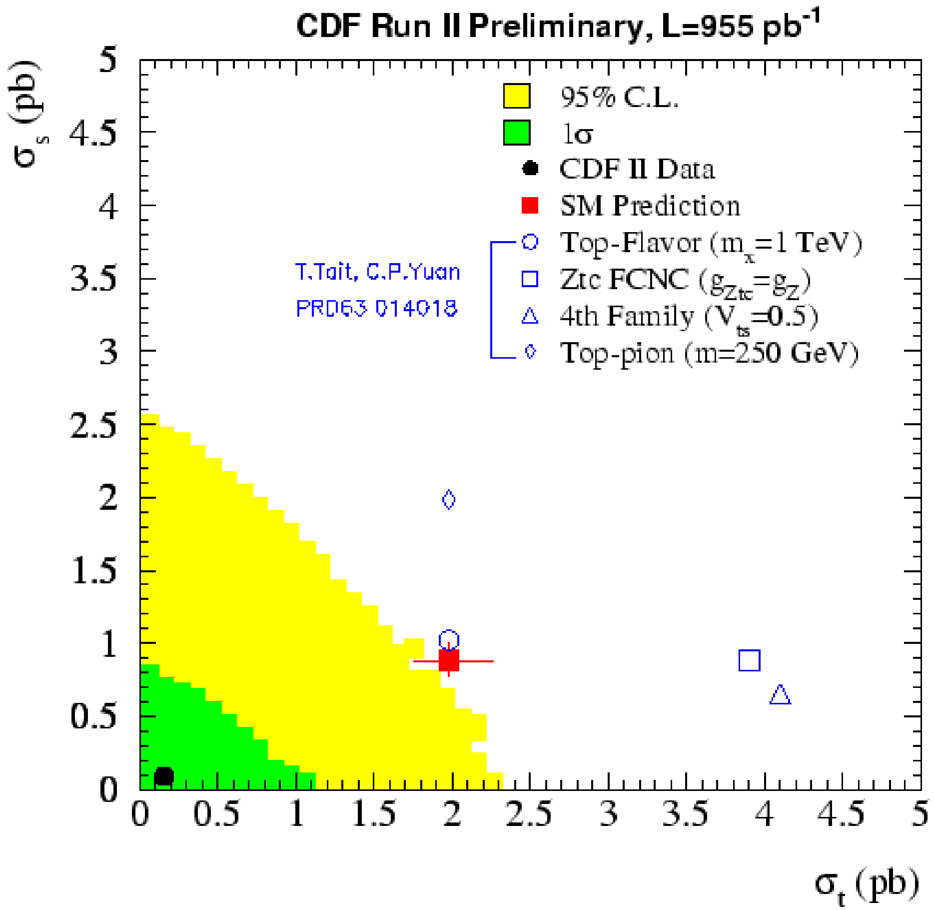
Best fit Combined search

$$\sigma_{s+t} = 0.8^{+1.3}_{-0.8}\,\text{pb}$$

at an expected sensitivity of 2.6σ

Result compatible with no single top and with SM single top.

Likelihood Analysis



Result excludes models beyond the SM

No excess over backgroud observed.

Best fit Separate Search:

$$\sigma_s = 0.1^{+0.7}_{-0.1}\,\text{pb}$$

$$\sigma_t = 0.2^{+0.9}_{-0.2}\,\text{pb}$$

Best fit Combined search

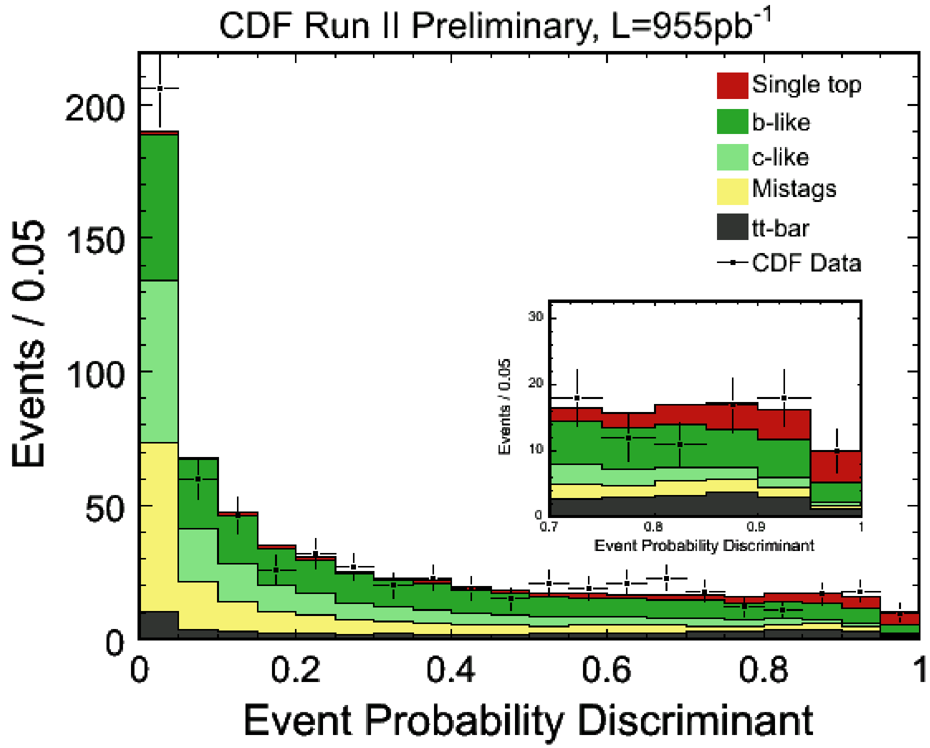
$$\sigma_{s+t} = 0.3^{+1.2}_{-0.3}\,\text{pb}$$

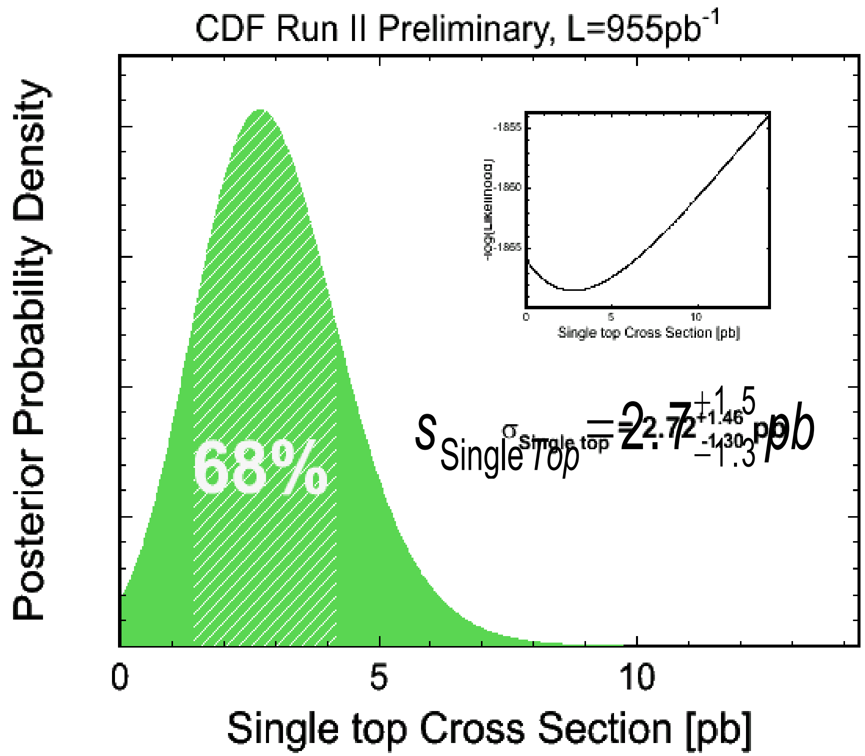
Limits (95%CL):

Expected: $\sigma_{s+t} < 2.9\,\text{pb}$

Observed: $\sigma_{s+t} < 2.7\,\text{pb}$

Matrix Element Analysis





- Matrix Element analysis observes excess over background expectation
- Likelihood fit result for combined search: $\sigma_{s+t} = 2.7^{+1.5}_{-1.3}\,\text{pb}$

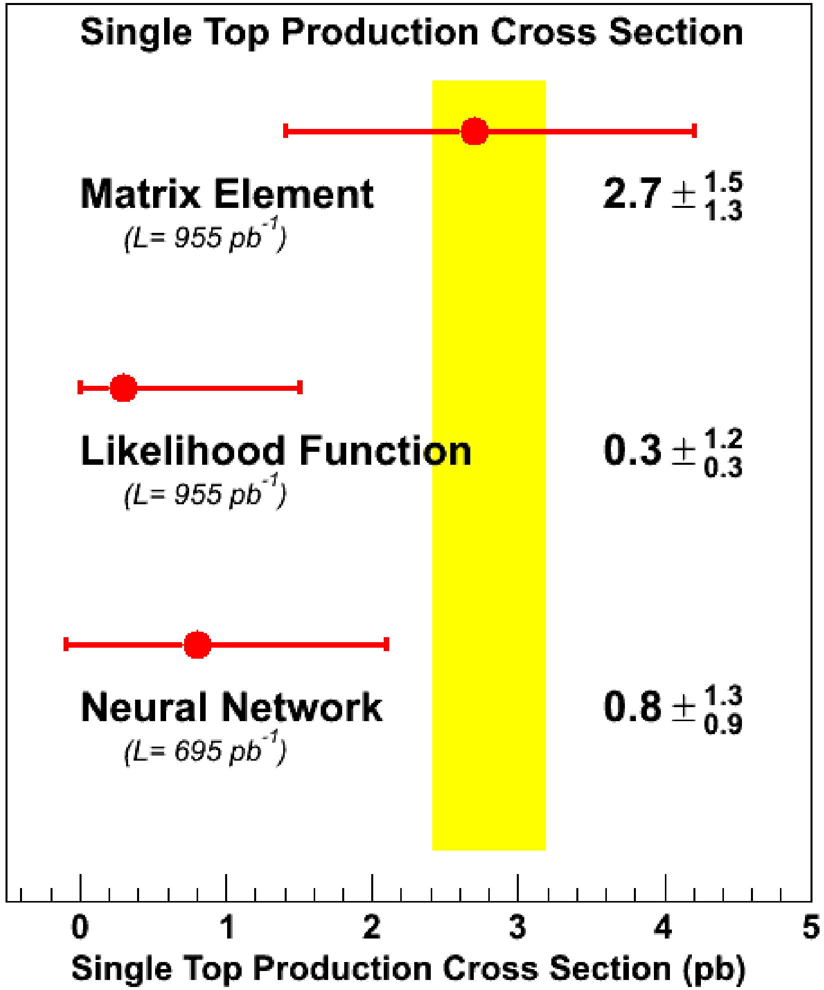
expected sensitivity: 2.5σ

Compatibility of CDF Results

Performed common pseudo-experiments

- Use identical events
- ME uses only 4-vectors of lepton, Jet1/Jet2
- LF uses sensitive event variables
- Correlation among fit results: $\sim 53\%$

6% of the pseudo-experiments had a difference in fit results at least as large as the difference observed in data



Measurement of $|V_{tb}|$

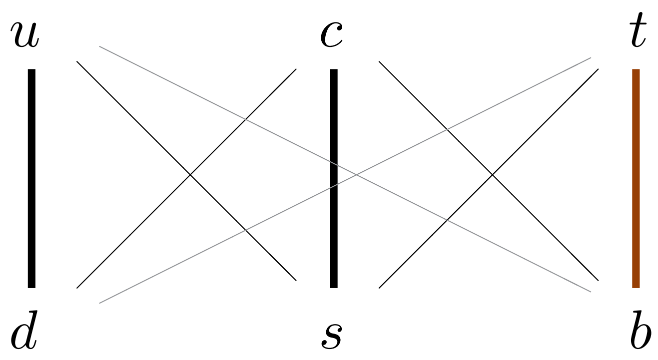
Weak Force Mixing

Three families of matter

e	electron	μ	muon	τ	tauon
ν_e	electron neutrino	ν_μ	muon neutrino	ν_τ	tau neutrino
u	up-quark	c	charm-quark	t	top-quark
d	down-quark	s	strange-quark	b	bottom-quark

Conversion of Quark Types

Only the $W^\pm$ bosons observed to convert quark types.



Conversions

- within one family most frequent
- between adjacent families more seldom
- between 1st and 3rd family very rare

Standard Model Picture Cabbibo-Kobayashi-Maskawa Matrix

Observable remainder of differences in weak and mass eigenstates:

$$V_{\mathrm{CKM}} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}$$

Unitary: $VV^{+} = 1$ ("number of ingoing quarks equal number of outgoing quarks")

Existing measurements and unitarity in a 3 quark family model yield: $V_{tb} \geq 0.998$

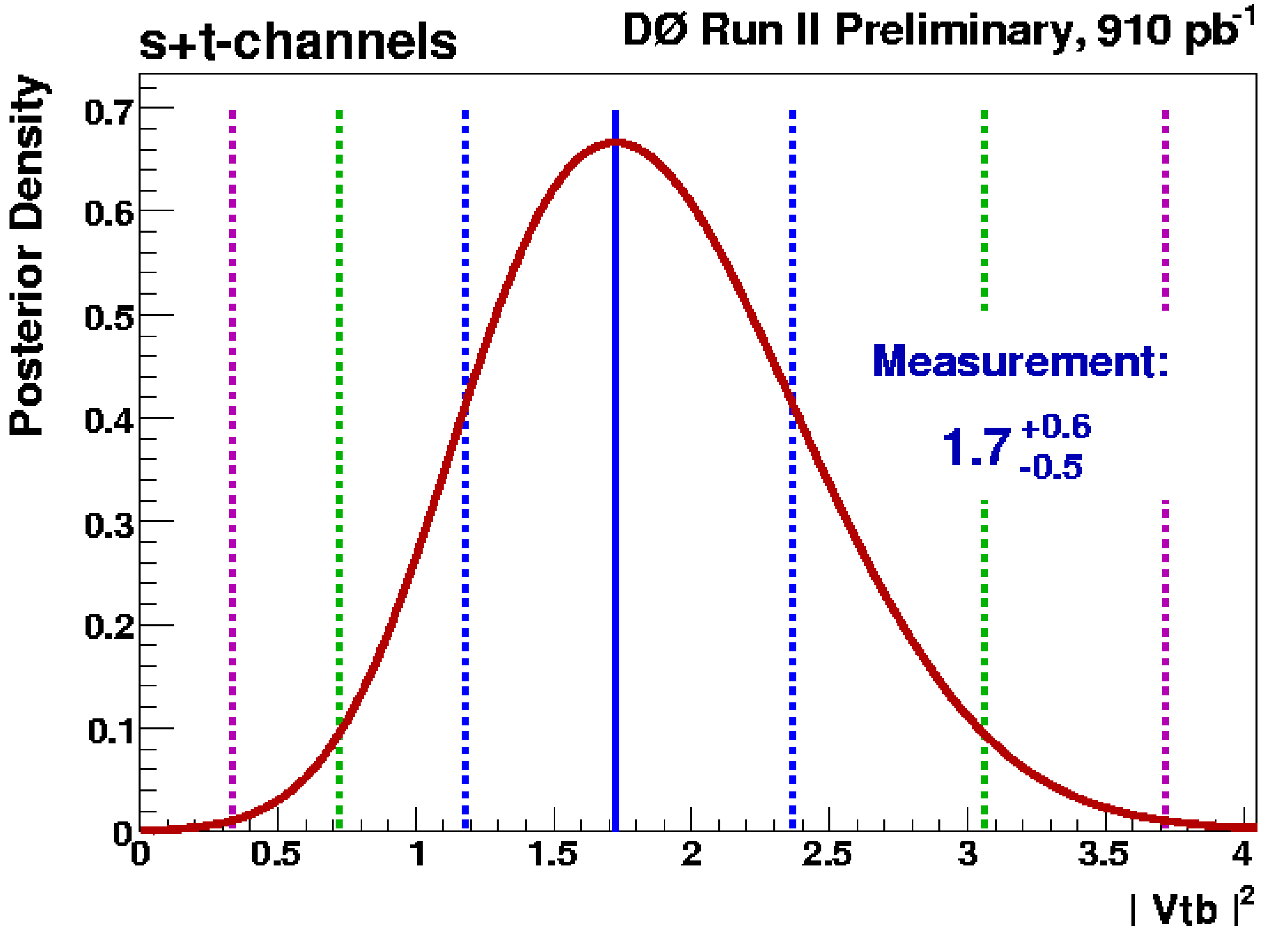
But additional quark families would reduce constraints.

Measuring $|V_{tb}|$

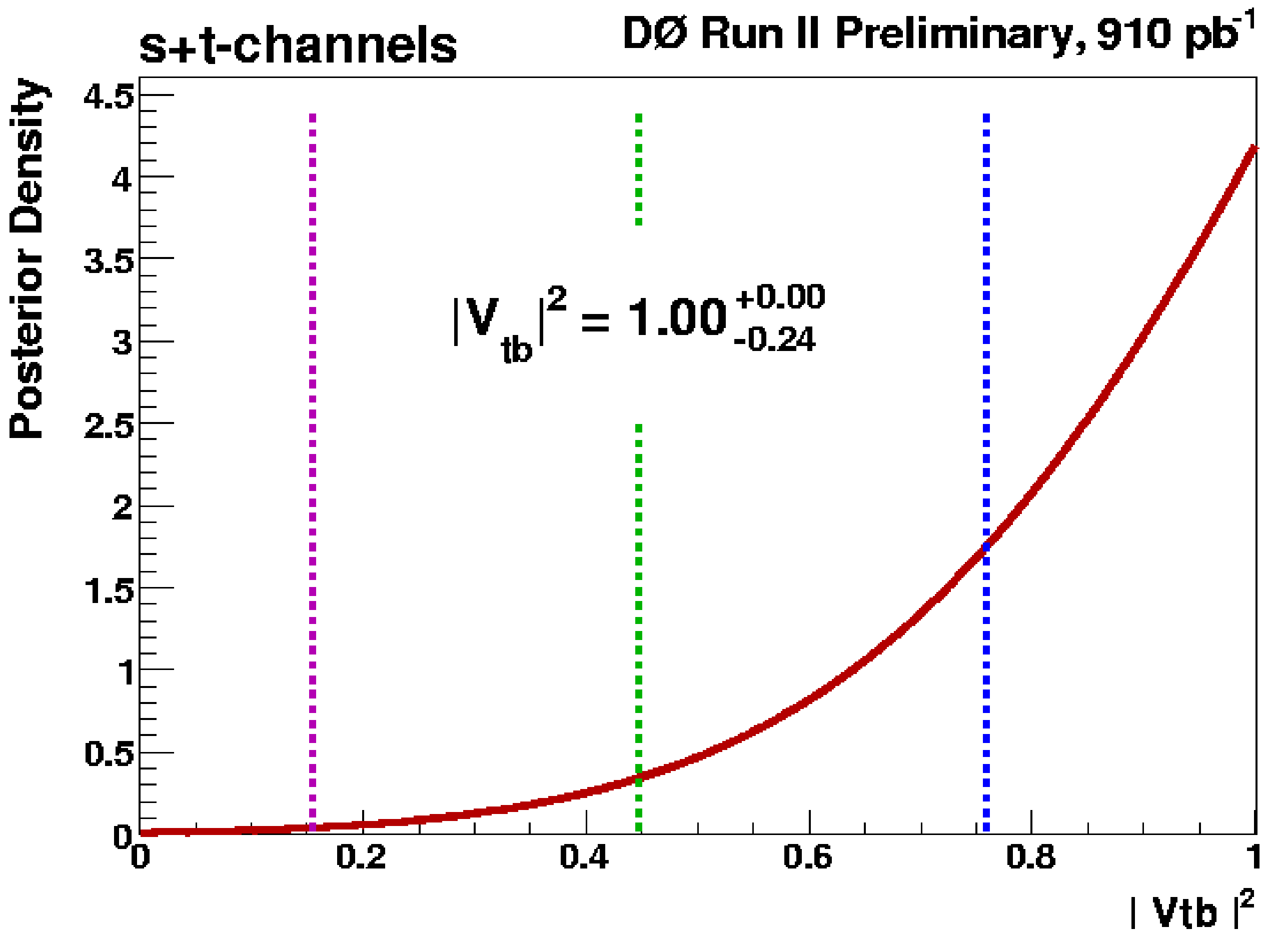
Use same techniques as cross section measurement, but make a posterior in $|V_{tb}|^2$.

Caveat: assume SM top quark decays.

Posterior Density



Result for $|V_{tb}|^2$



Constrain $|V_{tb}|$ to physical region and integrate: $|V_{tb}| = 1.00^{+0}_{-0.12}$

Summary

Summary

- Weak single top production is an interesting channel to challenge the SM.

- DØ has found first evidence for single top production at the Tevatron:
  - for sum of s- and t-channel combined analysis yields: $\sigma_{s+t} = 4.7 \pm 1.3\,\mathrm{pb}$
  - a 3.6σ effect above 0-signal hypothesis.
  - the 3 alternative analyses give consistent results

- CDF results are inconclusive (yet).
  - Despite better sensitivity no evidence found.
  - CDF got unlucky.

- Both experiments are consistent with SM expectation.

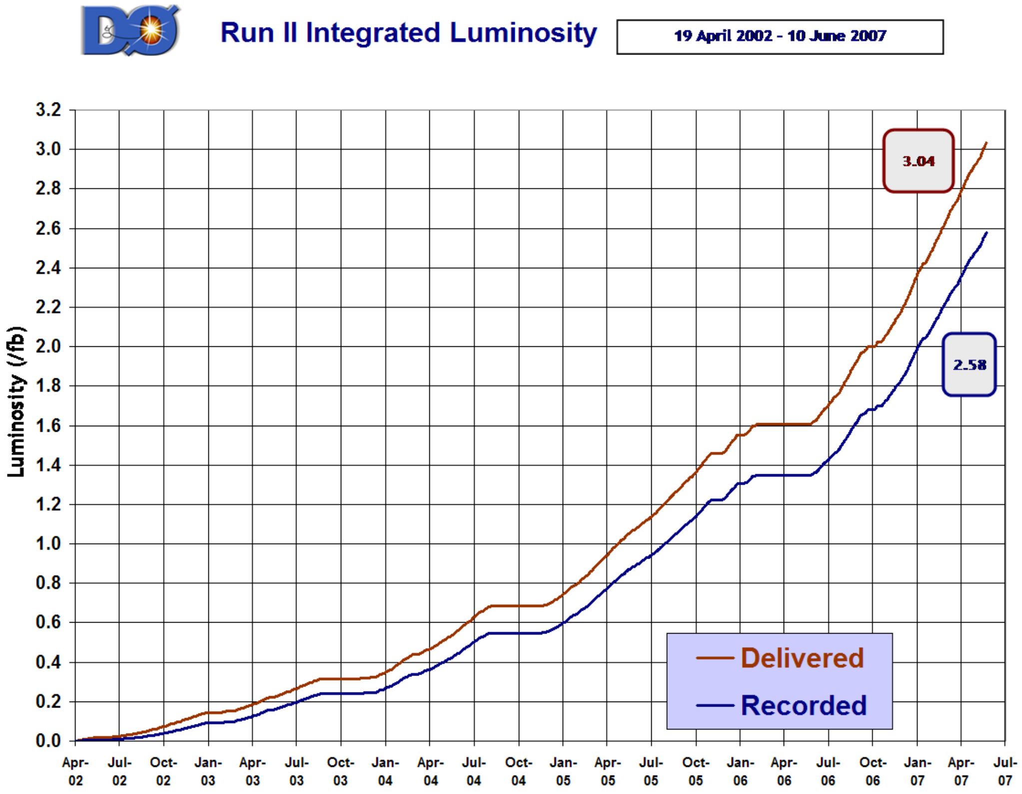
- DØ presented a first direct measurement of V_{tb} yields: $|V_{tb}| = 1.00^{+0}_{-0.12}$

Outlook

- Both experiments already have more data at hand and collecting.

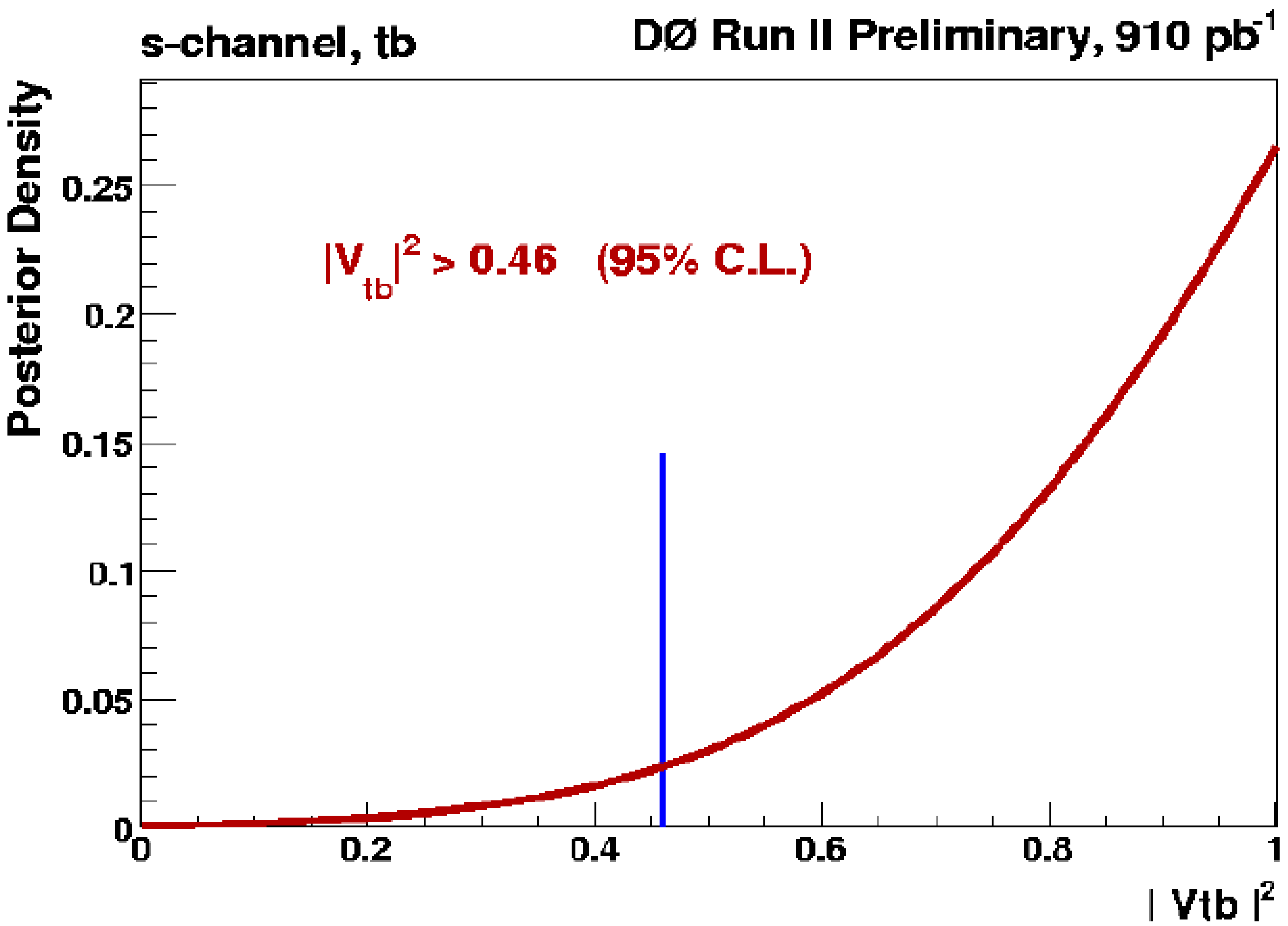
- LHC will bring
  way more statistics.

- Both measurements will be
  systematics dominated.



Backup

DØ V_{tb} Limit


s-channel, tb
DØ Run II Preliminary, 910 pb-1
Posterior Density
|Vtb|2 > 0.46 (95% C.L.)
0.25
0.2
0.15
0.1
0.05
0
0
0.2
0.4
0.6
0.8
1
| Vtb |2

Systematic Uncertainties

Relative Systematic Uncertainties			
$t\bar{t}$ cross section	18%	Primary vertex	3%
Luminosity	6%	Electron reco * ID	2%
Electron trigger	3%	Electron trackmatch & likelihood	5%
Muon trigger	6%	Muon reco * ID	7%
Jet energy scale	wide range	Muon trackmatch & isolation	2%
Jet efficiency	2%	$\varepsilon_{\mathrm{real}-e}$	2%
Jet fragmentation	5–7%	$\varepsilon_{\mathrm{real}-\mu}$	2%
Heavy flavor fraction	30%	$\varepsilon_{\mathrm{fake}-e}$	3–40%
Tag-rate functions	2–16%	$\varepsilon_{\mathrm{fake}-\mu}$	2–15%